

```

Mouse      --MLVATLLCALCCGLLAASAHAGYSEDRCSWRGSGLTQEPGSGVGLTLDTCEGAIEWLY 58
Rat        --MLVAAALLCALCCGLLAASARAGYSEDRCSWRGSGLTQEPGSGVGLTLDTCEGAIEWLY 58
Human      MGFPAAALLCALCCGLLAPAAARAGYSEERCSSWRGSGLTQEPGSGVGLALACAEGAVEWLY 60
           : .*:*****. *:*****.*****: * *:***:***

Mouse      PAGALRLTLGGPDPGTRPSIVCLRPERPFAGAQVFAERMTGNLELLLLAEGPDLAGGRCMR 118
Rat        PAGALRLTLGGSDPGTRPSIVCLRPTRFAGAQVFAERMAGNLELLLLAEGQLAGGRCMR 118
Human      PAGALRLTLGGPDPRARPGIACLRPVRPFAGAQVFAERAGGALLELLAEGPGPAGGRCVR 120
           *****.* *:*.***** ***** * ***** .*****:*

Mouse      WGPERRRALFLQATPHRDISRRVAAFRFLHEDQRAEMSPQAQGLGVDGACRPCSDAELL 178
Rat        WGPERRRALFLQATPHRDISRRVAAFQFELHEDQRAEMSPQAQGFVGVDGACRPCSDAELL 178
Human      WGPERRRALFLQATPHQDISRRVAAFRFLREDGRPELPPQAHGLGVDGACRPCSDAELL 180
           *****:*****:***:* *.:***:*:*****:*****

Mouse      LAACTSDFVIHGTHIGVAHDTELQESVITVVVARVIRQTLPLFKESGSSEQGRASIRTLL 238
Rat        LTACTSDFVIHGTHIGVVHDMELQESVITVVATRIVIRQTLPLFQESSEGRQASVRTLL 238
Human      LAACTSDFVIHGTHIGVTHDVELQESVITVVAARVLRQTPPLFQAGRSGDQGLTSIRTPL 240
           *:***** *****.* *****.:*:*** ***: * *.:* :*:**

Mouse      RCGVPRPGPSFLFMGWSRFGEAWLGCAPRFQEF SRVYSAALTTHLNPCEMALD 291 (SEQ ID NO: 26)
Rat        RCGVPRPGPSFLFMGWSRFGEAWLGCAPRFQEF SRVYSAALAAHLNPCEVALD 291 (SEQ ID NO: 13)
Human      RCGVHPGPGTFLFMGWSRFGEARLGCAPRFQEF RRAYEAARAAHLHPCEVALH 293 (SEQ ID NO: 3)
           *****:*****:***** * * *..*****

```

scoring matrix: BLOSUM50, gap penalties: -12/-2  
 42.3% identity; Global alignment score: 747

```

      10      20      30      40      50 *
Innog. MRGAARAAWGRAGQPWPRPPAPGPPPPPLPLLLLLLAGLLGGAG-AQYSSDRCSWKGSGL
      :      :      :      :      :      :      :      :      :
NsG33  -----MGFPAAALLCALCCGLLAPAARAGYSEERCSSWRGSGL
              10      20      30

      60      70      *  80      90      100      *110
Innog. THEAHRKEVEQVYLRCAAGAVEWMYPTGALIVNLR-PNTFSPARHLTVCI RSFTDSSGAN
      :      :      :      :      :      :      :      :      :
NsG33  TQEPGS--VGQLALACAEGAVEWLYPAGALRLTLGGPDPR--ARPGIACLRPVRPFAGAQ
      40      50      60      70      80      90

      120     130     140 *   150     160     170
Innog. IYLEKTG-ELRLLVPDGDGRPGRVQC--FG-LEQGGLFVEATPQDIGRRTTG FQYELVR
      :      :      :      :      :      :      :      :      :
NsG33  VFAERAGGALELLLAEGPG-PAGGRCVRWGPRERRALFLQATPHQDISRRVA AFRFELRE
      100     110     120     130     140     150

      180     190     200     210     220
Innog. RHRAS---DLHEL SAP--CRPCSDTEVLLAVCTSDFAVRGSIQQVTHEPERQDS AIHLRV
      :      :      :      :      :      :      :      :      :
NsG33  DGRPELPPQAHGLGVDGACRPCSDAELLAACTSDFVIHGIIHGVT HDVELQESVITVVA
      160     170*   *   180   *   190     200     210

      230     240     250     260     270     280
Innog. SRLYRQKSRVFEPVPEGDGHWQG--RVRTLLECGVRPGHGD FLTG HMFGEARLGCAPR
      :      :      :      :      :      :      :      :      :
NsG33  ARVLRQTPLPFQAGRSGD---QGLTSIRTPLRCGVHPGPGTFLFMGWSRFGEARLGCAPR
      220     230     240 *   250     260     *

      290     300     * 310
Innog. FKDFQRM YRDAQERGLNPCEVGT D (SEQ ID NO: 40)
      :      :      :      :
NsG33  FQEFRRAYEAARAAHLHPCEVALH (SEQ ID NO: 3)
      270     280     290
  
```

Fig 5

6/13

```
1  gcttcgcgg ggccggggcg cgccggcccc cggctgtcc cggccgcgcc cggaccgcg cccgcgcgg gcagcgtgg tgagagcccc gactccccg
101 acgcgcgcc cgtgtccatg gggttcccc cgcgggcgt gctctgcg cgtgtgtcg gctcctggc cccggctgcc cgcgcgggt actcgcagga
    >>.....CDS.....>
    m g f p a a a l l c a l c c g l l a p a a r a g y s e
201 gcgctgacg tggaggggca cggcctcac ccaggagccc ggcagcgtgg ggcagctggc cctggcctgt gcggaggcg cgttgagtg gctgtaccg
    >.....CDS.....>
    e r c s w r g s g l t q e p g s v g q l a l a c a e g a v e w l y p
301 gctggggcg tgcgctgac cctgggcggc ccgatccca gagcgcgcc cggcatgcc tgtctgcgc cggctggcc cctcggggc gccaggtct
    >.....CDS.....>
    a g a l r l t l g g p d p r a r p g i a c l r p v r p f a g a q v
401 tcgcggagcg cgcagggggc gccctggagc tgctgtggc cgagggccc ggcgcggcag gggccgcgtg cgtgcgtgg ggtcccccg agcgcgggc
    >.....CDS.....>
    f a e r a g g a l e l l l a e g p g p a g g r c' v r w g p r e r r
501 cctcttctg cagggcacg cgcaccagga catcagccg cgcgtggcg ccttcgctt tgagctgcg gaggaaggc gccccgagt gccccgcag
    >.....CDS.....>
    a l f l q a t p h q d i s r r v a a f r f e l r e d g r p e l p p q
601 gccacggtc tcggcgtaga cgtgtcctgc aggccctgca gcgacgtga gctgtctctg gccgcatgca ccagcgactt cgtaattcac gggatcatcc
    >.....CDS.....>
    a h g l g v d g a c r p c s d a e l l l a a c t s d f v i h g i i
701 atggggtcac ccattgacgt gagctgcagg agtgtgtcat cactgtgtg gccgcccgtg tcctccgca gacaccgcg ctgttccagg cggggcgatc
    >.....CDS.....>
    h g v t h d v e l q e s v i t v v a a r v l r q t p p l f q a g r
801 cggggaccag gggctgacct ccattcgtac cccactgcg tgtggcgtcc accggggccc aggcacctc ctcttcattg gctggagccg ctttggggag
    >.....CDS.....>
    s g d q g l t s i r t p l r c g v h p g p g t f l f m g w s r f g e
901 gcccggtgg gctgtgccc acgattccag gatttccgc gtgcctacc ggtgcccgt gctgccacc tccaccctg cgagggtggc ctgcactgag
    >.....CDS.....>
    a r l g c a p r f q e f r r a y e a a r a a h l h p c e v a l h -
1001 gggctgggtg ctgggggagg gctggtagga gggagggtg gccactgct ttggagggtg tgggactatc aataagaact ctgttccagc aaaaaaaaa
1101 aaaaaaaaa (SEQ ID NO: 2)
```

(SEQ ID NO: 3)

Fig 6

7/13

```
1 ccacgcgtcc gccacgcgt ccgcgcttct ttgcgcgtc tgttgccgc tcttgccgc gtcgcgtcac gctggctact cggaagaaccg ctgcagctgg
>>.....CDS.....>
101 h a s a h a s a l l c a l c c g l l a a s a h a g y s e d r c s w
aggggcagcg gtttgaccca ggagcctggc agcgtgggc agctgacct ggactgtact gagggcgcta tcgagtggct gtaccacgct ggggcgctgc
>>.....CDS.....>
201 r g s g l t q e p g s v g q l t l d c t e g a i e w l y p a g a l
gcctgacct gggcgccccc gatccgggca cagggccccc catcgtctgt ctgcgccag agcggccctt cgctggcgcc caggtctctg ctgaacgtat
>>.....CDS.....>
301 r l t l g g p d p g t r p s i v c l r p e r p f a g a q v f a e r
gaccggcaat ctagagttgc tactggccga gggcccgac ctggctgggg gcgctgcat gcgctgggt cccgcgagc gccgagccct ttctctgcag
>>.....CDS.....>
401 m t g n l e l l l a e g p d l a g g r c m r w g p r e r a l f l q
gccacaccac accgcgacat cagcgcaga gttgctgcct tccgttttga actgcacgag gaccaacgtg cagaaatgc tccccaggct caaggtcttg
>>.....CDS.....>
501 a t p h r d i s r r v a a f r f e l h e d q r a e m s p q a q g l
gtgtggatgg tgcctgcagg cctgcagg atgcccagct cctcctggct gcctgcacca gtgattttgt gatccacggg accatccatg gggtcgcccc
>>.....CDS.....>
601 g v d g a c r p c s d a e l l l a a c t s d f v i h g t i h g v a
tgacacagag ctgcaagaat cagtcacac tbtgggtggt gctcgtgtca tccgccagac actgccactg ttcaagggaag ggagctcgga gggccaaggg
>>.....CDS.....>
701 h d t e l q e s v i t v v a r v i r q t l p l f k e g s s e g q g
cgggcctcca ttcgtacctt gctgcgctgt ggtgtgcgtc ctggcccagg ctctctctc ttcattggct ggagccgatt tggcgaagct tggctgggt
>>.....CDS.....>
801 r a s i r t l l r c g v r p g p g s f l f m g w s r f g e a w l g
gtgctcccc cttccaagag ttacgctgtg ttatttcagc tgctctcag acccatctca accatgtga gatggcactg gactgagaga cctgggagca
>>.....CDS.....>
901 c a p r f q e f s r v y s a a l t t h l n p c e m a l d - (SEQ ID NO: 8)
agccctggat ggaccttctt ctggagatgg ggtgttgggg aggtgatgg gaggtgggt gagaagggtg tggctcggat ggcattcctgg taccacagt
1001 gagctggtag aatactaagt aatctggacc ataaaaaaa aaaaaaaa (SEQ ID NO: 7)
```

Fig 7a

```
1  gggcagccgc gccgcgggct gctcgcgctg cggccccgc cctcccggg cagcagtcgc agggccccgc gcgtcccta accatgctgg
   >>CDS.>
   m l
91  tagccacgct tctttgcgc cctgtgttgc gctcctggc cgcgtcgcct cagctggct actcggaaga ccgctgcgc tggaggggca
   >.....CDS.....
   v a t l l c a l c c g l l a a s a h a g y s e d r c s w r g
181  gcggtttgac ccaggagcct ggcagcgttg ggcagctgac cctggactgt actgaggcg cttacgagt gctgtacca gctggggcgc
   >.....CDS.....
   s g l t q e p g s v g q l t l d c t e g a i e w l y p a g a
271  tgcgcctgac cctggggcgc ccgataccg gcaacgcgc cagcactgc tgtctggcc cagagcgcc cttcgtggt gccaggtct
   >.....CDS.....
   l r l t l g g p d p g t r p s i v c l r p e r p f a g a q v
361  tcgctgaacg tatgacgcg aatctagagt tgcactggc cgaggcccg gacctggctg gggccgctg catgcgctg ggtcccccgc
   >.....CDS.....
   f a e r m t g n l e l l l a e g p d l a g r c m r w g p r
451  agcgcgagc ccttttctg caggccacac cacaccgca catcagcgc agagtgtg cttccggtt tgaactgcac gaggaccaac
   >.....CDS.....
   e r r a l f l q a t p h r d i s r r v a a f r f e l h e d q
541  gtgcagaaat gtctcccag gctcaagtc ttggtgtgga tgggtcctgc aggccctgca gtgatgccga gctcctctg gctgcatgca
   >.....CDS.....
   r a e m s p q a q g l g v d g a c r p c s d a e l l l a a c
631  ccagtgttt tgtgatccac gggaccatcc atgggggtgc ccatgacaca gagctgcaag aatcagtcac cactgtggtg gttgctcgtg
   >.....CDS.....
   t s d f v i h g t i h g v a h d t e l q e s v i t v v a r
721  tcatccgcca gacactgcca ctgttcaagg aaggagctc ggagggccaa ggccggcct ccattcgtac cttgctgcgc tgtggtgtgc
   >.....CDS.....
   v i r q t l p l f k e g s s e g q g r a s i r t l l r c g v
811  gtctggccc aggtccttc ctcttcattg gctggagccg atttggcga gcttgctgc ccgttccaa gagttcagcc
   >.....CDS.....
   r p g p g s f l f m g w s r f g e a w l g c a p r f q e f s
901  gtgtctattc agctgcttc acgaccatc tcaacccatg tgagatggca ctggactgag agacctggga gcaagccctg gatggacatt
   >.....CDS.....
   r v y s a a l t t h l n p c e m a l d - (SEQ ID NO: 26)
991  cttctggaga tggggtgtg gggagggtga tgggagggtg ggtgagaagg gttggctcg gatgcatcc tggtaaccac agtgagctgg
1081  tagaatacta agtaaatctg accataccag ccactgtagt catggtcttc tgtggcaggc agcataccca gctctgtgcc tgcctcactt
1171  tgtctactct ccagctctgt gcccttctaa cccttcttag cctgctgacc agtgagctca tgtttcctc gaattccagg gtgctgctgg
1261  ggttcagagc aaccgtgcg tagtttgga gacttgagct aattgtttt tttttgtttg ttttttgtt tggttaagg tggcctgggg
1351  gggggcgcaa aca (SEQ ID NO: 25)
```

Fig 7b

9/13

```

1  atgctggttag cggcgcttct ctgcgcgctg tgctgcggcc tcttggtcgc gtccgctcga gctgggtact ccgaggacgc ctgcagctgg aggggcagcg
>>.....CDS.....>
    m l v a a l l c a l c c g l l a a s a r a g y s e d r c s w r g s
101 gtttgacca ggaacctggc agcgtgggc agctgacct ggattgtact gaggtgcta tccagtgct gatatcagct gggcgctgc gcctgactct
>>.....CDS.....>
    g l t q e p g s v g q l t l d c t e g a i e w l y p a g a l r l t
201 aggggctct gatccgggca cgcggcccg catcgtctgt ctgcgcceaa cagggccctt cgctgggtgc caggtcttcg ctgaacggat ggcgggaac
>>.....CDS.....>
    l g g s d p g t r p s i v c l r p t r p f a g a q v f a e r m a g n
301 cttaggtgc tactggcca gggccaaggc ctggctggg gccgtgcat gcgtggggt cctgcgcgc gccgagccct ttctctcag gccacgccac
>>.....CDS.....>
    l e l l l a e g q g l a g g r c m r w g p r e r a l f l q a t p
401 accgggacat cagccgcaga gttgtgcct tccaattga actgcacgag gaccaacgtg cagaaatgc tccccaggcc caaggttttg gtgtggatgg
>>.....CDS.....>
    h r d i s r r v a a f q f e l h e d q r a e m s p q a q g f g v d
501 tgcctgcagg ccttgcagt atgcgcgagt ccttctgact gcattgcacca gtgactttgt gatccatggg accatccatg gggctcgtcca tgacatggag
>>.....CDS.....>
    g a c r p c s d a e l l l t a c t s d f v i h g t i h g v v h d m e
601 ctgcaagaat cagtcacac tgtggtggcc actcgtgtca tccgcacgac actgccactg ttccaggaag ggagctcggg caggccctccg
>>.....CDS.....>
    l q e s v i t v v a t r v i r q t l p l f q e g s s e g r g q a s
701 ttcgtacctt gttgcgtgt ggtgtgcgtc ctggcccagg ctccttcctc ttcatggggt ggagccgatt tggcgaagct tggctgggt gcgctccccg
>>.....CDS.....>
    v r t l l r c g v r p g p g s f l f m g w s r f g e a w l g c a p
801 cttccaagag ttacgcgtg totattcagc tgctctcgc gccacactca acccatgtga ggtggcactg gactgagaga cctggggagca agccctggat
>>.....CDS.....>
    r f q e f s r v y s a a l a a h l n p c e v a l d - (SEQ ID NO:13)
901 ggatcttct ctgggggatgg ggtgttgggg aggggtgata ggagggtggg tggggaagggt gtggctcaga tggcatcctg gtaccacag tgagggtggta
1001 gaatactaaa taacctggat cacacc (SEQ ID NO:12)

```

Fig 8